

**OPTICAL CHARACTER RECOGNITION (OCR) FOR MOBILE
APPLICATION**

ANAS ABDELSATAR MOHAMMAD SALAMEH

UNIVERSITI UTARA MALAYSIA 2010

OPTICAL CHARACTER RECOGNITION (OCR) FOR MOBILE APPLICATION

**A project submitted to Faculty of Information Technology in partial
fulfillment of the requirements for the degree Master
(Information Technology),
Universiti Utara Malaysia**

By

Anas Abdelsatar Mohammad Salameh



KOLEJ SASTERA DAN SAINS
(College of Arts and Sciences)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certifies that)

ANAS ABDELSATAR MOHAMMAD SALAMEH
(806151)

calon untuk Ijazah
(candidate for the degree of) **MSc. (Information Technology)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project of the following title)

OPTICAL CHARACTER RECOGNITION (OCR) FOR MOBILE APPLICATION

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan
dan meliputi bidang ilmu dengan memuaskan.
(that this project is in acceptable form and content, and that a satisfactory
knowledge of the field is covered by the project).

Nama Penyelia
(Name of Supervisor) : **ASSOC. PROF. ABDUL NASIR ZULKIFLI**

Tandatangan
(Signature) : Abd. Nasir Tarikh (Date) : 19/10/10

Nama Penilai
(Name of Evaluator) : **ASSOC. PROF. DR. ZULIKHA JAMALUDDIN**

Tandatangan
(Signature) : [Signature] Tarikh (Date) : 16/11/10

PERMISSION TO USE

In presenting this project in partial fulfillment of the requirements for a postgraduate degree from the Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this project in any manner in whole or in part, for scholarly purposes may be granted by my supervisor(s) or in their absence by the Dean of Postgraduate Studies and Research. It is understood that any copying or publication or use of this project or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my project.

Requests for permission to copy or to make other use of materials in this project, in whole or in part, should be addressed to

**Dean of Postgraduate Studies and Research
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman
Malaysia**

ABSTRACT

The usefulness of integrating different techniques in wireless applications has brought the needs for providing better services in different technical sectors. Wireless Application Protocol (WAP) has been widely used for obtaining the required connection between clients via their handheld devices. This study highlights the difficulties that are faced by travelers in understanding foreign text during their journeys to other countries with different native languages. Hence, this study aimed to provide a solution by developing a mobile application based optical character recognition (OCR) for extracting the textual elements from the images. Asprise used in this study to extract the image text contents, meanwhile, Google API translation also used to translate the extracted contents into the selected language. The experiment result indicated that using Asprise OCR in extracting the text elements from the image was high accuracy among the free and simple OCR.

ACKNOWLEDGEMENTS

Praise to Allah for his guidance and blessing for giving me the strength and perseverance to complete this project. I would foremost like to thank my parents, for providing me with the opportunity to pursue my goals and for their love and affection, which has helped me through the most trying times. Equal gratitude goes out to my siblings and brothers. I would like to thank my supervisor: **Assoc. Prof. Abdul Nasir Zulkifli** for his guidance and constant motivation that has enabled me to complete my project work. Moreover, I would also like to thank him for the opportunities that he has made available to me.

Anas Abdelsatar Mohammad Salameh

LIST OF TABLES

Table 2.1	Hardware, Scanning, and Software Requirements	13
Table 2.2	Comparison between OCR applications	16
Table 2.3	Comparison between online translations	18
Table 4.1	List of software	34
Table 5.1	Extracted Test Result based on using the adopted OCR	43
Table 5.2	Extracted Test Result based on using the Free OCR	45
Table 5.3	Extracted Test Result based on using Simple OCR	46

LIST OF FIGURES

Figure 1.1	Mobile Translator	5
Figure 2.1	The General Process of Optical Character Recognition (OCR)	10
Figure 2.2	OCR elements	11
Figure 2.3	Document scanning process	13
Figure 2.4	Recognition process	14
Figure 2.5	Verifying process of OCR	14
Figure 2.6	Image Processing for Mobile Devices	20
Figure 2.7	Client and Server architecture for mobile image recognition	20
Figure 2.8	Image recognition algorithm	21
Figure 2.9	Text Translation Markup Language (TTML) process	22
Figure 2.10	The spoken dialogue translation system between two clients	23
Figure 2.11	Virtual Data Translation	24
Figure 2.12	Indian Dictionary System	25
Figure 3.1	General methodology	26
Figure 3.2	The recognition process of the proposed application	28
Figure 3.3	Image translation process	29
Figure 4.1	Mobile Translation Use Case Diagram	36
Figure 4.2	GUI for translating from Malay to English	37
Figure 4.3	GUI result for translating from Malay to English	38
Figure 4.4	GUI for translating from Malay to Arabic	38
Figure 4.5	GUI result for translating from Malay to Arabic	39
Figure 5.1	Extracted Test Result based on using the adopted OCR	44
Figure 5.2	Extracted Test Result based on using the Free OCR	45
Figure 5.3	Extracted Test Result based on using the Simple OCR	47

TABLE OF CONTENTS

PERMISSION TO USE	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER ONE: INTRODUCTION	
1.0 Introduction	1
1.1 Research Problem	3
1.2 Research Questions	4
1.3 Research Objectives	4
1.4 Research Scope	4
1.5 Research Significance	5
1.6 Research Structure	6
1.7 Summary	7
CHAPTER TWO: LITERATURE REVIEW	
2.0 Introduction	8
2.1 Optical Character Recognition	10
2.1.1 Function of OCR	12
2.1.2 Features of OCR	12
2.1.3 OCR Requirements	13
2.1.4 OCR Process Stages	13
2.1.4.1 Document Scanning Process	14
2.1.4.1 Recognizing process	14
2.1.4.1 Verifying Process	15
2.2 OCR for Mobile Device	15
2.5 Comparison between OCR applications	16
2.5 Translation	18
2.6 Comparison between Online Translations	19
2.6 Mobile translator based on OCR	20
2.7 Related Works	21

CHAPTER THREE: RESEARCH METHODOLOGY

3.0 Introduction	28
3.1 General Methodology	29
3.1.1 Awareness of Problem	29
3.1.2 Suggestion	30
3.1.3 Development	31
3.1.4 Evaluation	32
3.1.5 Conclusion	33
3.2 Summary	33

CHAPTER FOUR: MOBILE TRANSLATION DESIGN

4.0 Introduction	34
4.1 Requirements Gathering	34
4.1.1 Structuring Mobile Translation Requirements	35
4.1.2 Mobile Translation Functional Requirements	36
4.1.3 Mobile Translation Software Requirements	36
4.1.4 Non Functional Requirements	37
4.1.4.1 Usability	37
4.1.4.2 Understandability	37
4.1.4.3 Reliability	38
4.1.4.4 Availability	38
4.2 Mobile Translation Use Case Diagram	38
4.3 User Interface of Mobile Translation	39
4.3.1 Header	40
4.3.2 Cover display	40
4.3.3 Headlines	40
4.3.4 Menu	40
4.4 Content Design	42
4.4.1 Content display	43
4.4.2 The navigational structure	43
4.5 Summary	43

CHAPTER FIVE: EVALUATION

5.1 Introduction	44
5.2 Summary	50

CHAPTER SIX: CONCLUSION

6.0 Introduction 52

6.1 Discussion..... 52

6.2 Future Studies 53

6.3 Conclusion 54

REFERENCES..... 55

APPENDIX..... 61

CHAPTER ONE

INTRODUCTION

1.0 Introduction

Wireless Application Protocol (WAP) is defined by International Engineering Consortium as an application environment and set of communication protocols for wireless devices to communicate with each other and with any external application (Yi, Cerone, & Zhang, 2006). The purpose of WAP was designed to facilitate the manufacturers, vendors and technology users around the world to independently access to the Internet and advanced telephony services. WAP technology can eliminate the gap between the mobile devices and the Internet and then can provides for mobile services. Mobile devices are small handheld devices such as mobile phones, palmtop computers and devices which special operating system.

Mobile devices also include Personal Digital Assistants (PDAs) with or without networking capabilities and mobile phones that may or may not be able to access the web. Handheld computers are mobile, flexible devices that can provide real-time, one-to-one support for students from within the context of their learning. In general, the characteristics of mobile devices convergence of mobile, handheld and wireless communication technologies as the newest technological revolution.

The contents of
the thesis is for
internal user
only

REFERENCES

- Agrawal, A., & Tandon, R. (2009). Optical Character Recognition-Shape Invariant Approach and Implications of Spatial Resolution. *Computer Vision and Image Processing*, 5(11) 1-6.
- Bazzi, I., Schwartz, R., & Makhoul, J. (1999). An Omnifont Open-Vocabulary OCR System for English and Arabic. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(6), 495-504.
- Chellapilla, K., Simard, P., & Nickolov, R. (2005). Fast Optical Character Recognition through Glyph Hashing for Document Conversion. *computer Society*, 2(15), 1-5.
- Dennis, A., Wixom, B., & Tegarden, D. (2005). *Systems Analysis with UML Version 2.0* (2nd ed.): New York. John Wiley & Sons.
- Doermann, D., Liang, J., & Li, H. (2003). Progress in Camera-Based Document Image Analysis. *Document Analysis and Recognition*, 1(8), 606.
- Flores, G. (2006). Language barriers to health care in the United States. *New England journal of medicine*, 355(3), 229-231.
- Fonseca, M., Barroso, B., Ribeiro, P., & Jorge, J. (2004). Retrieving Clipart Images by Content. *Image and Video Retrieval*, 3115(12), 500-507.
- Gonzalez, R., Woods, R., & Eddins, S. (2004). *Digital Image Processing Using MATLAB* (2nd ed.) New Jersey: Prentice Hall Upper Saddle River.
- Gouadec, D. (2007). *Translation as A Profession* (1st ed.). Netherlands: John Benjamins Publishing Co.
- Harris, R., Bala, P., Songan, P., Lien, E., & Trang, T. (2001). Challenges and Opportunities in Introducing Information and Communication Technologies to the Kelabit Community of North Central Borneo. *New Media & Society*, 3(3), 270-295.
- Hoffer, J., George, J., & Valacich, J. (1999). *Modern Systems Analysis and Design* (3rd ed.). London: Addison Wesley Longman.

- Hoffer, J., George, J., & Valacich, J. (2002). *Modern Systems Analysis and Design* (5th ed.). New Jersey: Prentice Hall Upper Saddle River.
- Hull, J., Liu, X., Erol, B., Graham, J., & Moraleda, J. (2010). *Mobile Image Recognition: Architectures and Tradeoffs*. Paper presented at the *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications*. (pp. 116-134). Annapolis, Maryland: IEEE xplore.
- Husain, S., Sajjad, A., & Anwar, F. (2007). *Online Urdu Character Recognition System*. Paper presented at the *MVA2007 IAPR Conference on Machine Vision Applications*. (pp. 84-96). Tokyo, JAPAN: Proquest.
- Jagannathan, L., & Jawahar, C. (2005). *Crosslingual Access of Textual Information Using Camera Phones*. Paper presented at the *Proceedings of the International Conference on Cognition and Recognition*. (pp. 18-33). Mandya, Mysore, Karnataka, India: IEEE xplore.
- Jagannathan, L., & Jawahar, C. (2005). *Perspective Correction Methods for Camera Based Document Analysis*. Paper presented at the *Proceedings of the First International Workshop on Camera-based Document Analysis and Recognition (CBDAR)*. (pp. 65-78). Seoul, Korea: IEEE-CS.
- Kimia, B. (2006). *Shape Representation for Image Retrieval, Image Databases*. New York: John Wiley & Sons.
- Kray, C., & Baus, J. (2003). *A Survey of Mobile Guides*. Paper presented at the *Workshop on HCI in mobile guides*, 5th Int. Symposium on HCI with Mobile Devices and Services. (pp. 133-150). Udine, Italy: ACM
- Liu, X., & Doermann, D. (2008). Mobile retriever: Access to Digital Documents from Their Physical Source. *International Journal on Document Analysis and Recognition*, 11(1), 19-27.
- Lucas, S., Panaretos, A., Sosa, L., Tang, A., Wong, S., & Young, R. (2003). ICDAR 2003 Robust Reading Competitions. *Document Analysis and Recognition*, 2(33), 682-687.
- Mirmehdi, M., Clark, P., & Lam, J. (2001). *Extracting low resolution text with an active camera for OCR*. Paper presented at the *Spanish Symposium on Pattern Recognition and Image Processing IX*. (pp. 55-72). Barcelona, Spain: IEEE

- Uwe Muegge. *Translation Contract: A Standards-based Model Solution*. Bloomington, India: AuthorHouse, 2005.
- Nakai, T., Kise, K., & Iwamura, M. (2006). Use of Affine Invariants in Locally Likely Arrangement Hashing for Camera-Based Document Image Retrieval. *Document Analysis Systems VII*, 3872(53), 541-552.
- Norbayah, M., & Norazah, M. (2007). Mobile Phone Usage for M-Learning: Comparing Heavy and Light Mobile Phone Users. *Campus-Wide Information Systems*, 24(5), 355-365.
- Paul, M., Okuma, H., Yamamoto, H., Sumita, E., Matsuda, S., Shimizu, T., & Nakamura, S. (2008). *Multilingual Mobile-Phone Translation Services for World Travelers. Paper presented at the 22nd International Conference on Computational Linguistics: Demonstration Papers*. (pp. 303-322). Ohio, USA: ACM.
- Philbin, J., Chum, O., Isard, M., Sivic, J., & Zisserman, A. (2007). *Object Retrieval with Large Vocabularies and Fast Spatial Matching. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition*. (pp. 1-8). San Francisco, California, USA: IEEE computer society.
- Rayner, M., Hockey, B., & Bouillon, P. (2006). Putting Linguistics into Speech Recognition: The Regulus Grammar Compiler. *Computational Linguistics*, 33(2), 300-327.
- Razzak, M., Hussain, S., Sher, M., & Khan, Z. (2009). *Combining Offline and Online Preprocessing for Online Urdu Character Recognition. Proceedings of the International MultiConference of Engineers and Computer Scientists*. (pp. 18-20). Bangkok, Thailand: ACM.
- Sekaran, U. (2003). *Research Methods For Business: A Skill Building Approach* (4th ed.). New Jersey: John Wiley & Sons, Inc.
- Shafait, F., Ul-Hasan, A., Keysers, D., & Breuel, T. (2006). *Layout Analysis of Urdu Document Images. Paper presented at the IEEE Multitopic Conference*. (pp. 207-225). Jakarta, Indonesia: INMIC '06. IEEE.

- Shanmugham, S., Monaco, P., & Eberman, B. (2005). A Media Resource Control Protocol (MRCP) Developed By Cisco, Nuance, and Speechworks. *Internet Engineering Task Force*.
- Starlander, M., & Estrella, P. (2009). *Relating Recognition and Translation Quality with Usability of Two Different Versions of MedSLT*. Proceedings of Machine Translation Summit XII. (pp. 301-305). Ottawa, Ontario, Canada: IEEE
- Steven, Z., Syed, O., & Stefan, W. (2007). Open Source OCR Framework Using Mobile Devices. *Interactive Multimedia Lab, Department of Electrical and Computer Engineering, Kent Ridge Crescent, Singapore*, 32(5), 1-6.
- Su, G., & Nieh, J. (2002). Mobile Communication with Virtual Network Address Translation. *Department of Computer Science, Columbia University*, 21(7) 1-14.
- Takacs, G., Chandrasekhar, V., Gelfand, N., Xiong, Y., Chen, W., Bismpianni, T., Grzeszczuk, R., Pulli, K., & Girod, B. (2008). *Outdoors Augmented Reality on Mobile Phone Using Loxel-Based Visual Feature Organization*. Paper presented at the *Proceeding of the 1st ACM international conference on Multimedia information retrieval*. (pp. 1-8). Vancouver, Canada: ACM
- Tsourakis, N., Bouillon, P., & Rayner, M. (2008). Design Issues for a Bidirectional Mobile Medical Speech Translator. *Document Analysis Systems VII*, 15(4), 55-78.
- Wachenfeld, S., Klein, H., Fleischer, S., & Jiang, X. (2007). *Segmentation of Very Low Resolution Screen-Rendered Text*. Paper presented at the *Ninth International Conference on Document Analysis and Recognition*. 2, (pp. 33-40). Curitiba, Parana, Brazil: ICDAR.
- Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., & Schmalstieg, D. (2008). *Pose Tracking From Natural Features On Mobile Phones*. Paper presented at the *Proceedings of the 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. (pp. 1-10). Cambridge, UK: IEEE/ACM
- Yi, X., Cerone, P., & Zhang, Y. (2006). *Secure Electronic Voting For Mobile Communications*. Paper presented at the *IEEE 63rd Vehicular Technology Conference*. (pp. 82-93). Melbourne, Australia : IEEE.
- Zhang, T., & Yu, B. (2005). Boosting With Early Stopping: Convergence and Consistency. *Annals of Statistics*, 33(4), 1538-1579.

Zikmund, G. (2003). *Business Research Methods* (7th ed.). Ohio: South Western Educational Publishing.